

## CHAPTER 11

*There is no Exclusion Problem*

STEINVÖR THÖLL ARNADOTTIR &amp; TIM CRANE

**0. Introduction**

Many philosophers want to say both that everything is determined by the physical and subject to physical laws and principles, and that certain mental entities cannot be identified with any physical entities. The problem of mental causation is to make these two assumptions compatible with the causal efficacy of the mental. The concern is that this physicalist picture of the world leaves no space for the causal efficacy of anything non-physical. The physical, as it is sometimes said, *excludes* anything non-physical from doing causal work.

The general shape of the problem is not new. Leibniz famously argued that Descartes's conception of the relationship between mind and body had no place for mental causation. On Descartes's view, according to Leibniz, the mind can only affect the body by changing the 'direction of motion' of the body's 'animal spirits'. Descartes had held that in this way the total 'quantity of motion' was conserved in psychophysical interaction. But Leibniz claimed that what should be conserved in these interactions is not quantity of motion but (as we would now put it) quantity of momentum (mass times velocity). So the mind cannot alter the *direction* of motion of the animal spirits without altering the quantity of *momentum* in the physical world. The physical law that Leibniz took himself to have discovered excludes the mental from making a causal difference.<sup>1</sup>

Leibniz's objection to Descartes was based on his view of the nature of the physical world. Contemporary philosophers also see the problem of mental causation as arising from assumptions about the physical world (see Papineau 1990). Partly because of the need to accommodate mental causation given these assumptions, many philosophers have sought to find a more intimate connection between the mental and the physical, holding that mental entities are *determined* by or *constituted* by physical

---

<sup>1</sup> For Leibniz's views, see Leibniz 1695, 1696. For contemporary discussion, including of the question of whether Leibniz's had correctly interpreted Descartes, see Garber 1983, and Woolhouse 1985.

entities. But the problem of mental causation has not gone away. The dominant worry about mental causation in the last few decades is that the physicalist principles to which most contemporary philosophers subscribe still leave no space for mental causation.

First, it seems that anyone who holds that everything is determined by the physical will have to say that the mental has physical effects if it is to have any effects at all. Whatever else the determination thesis might involve, it involves at least the thesis of global supervenience; the thesis that any minimal physical duplicate of the actual world is a duplicate in every respect—a duplicate *'simpliciter'*. (Jackson 1998; Lewis 1986, 1994). In other words: fix all the physical facts, and you fix all the facts there are. But if no changes are possible without the appropriate changes in the physical, then it is very reasonable to think that the mental must bring about physical changes if it is to bring about any changes at all.<sup>2</sup>

Second, there is another widely accepted principle known as the causal closure of the physical or completeness of physics, which claims that any physical effect has sufficient physical causes.<sup>3</sup> What this means is that physical entities alone fix the occurrence of any physical effect (either deterministically, or by fixing the chance of the effect). As we might put it, once all the physical entities are in place, you don't need to *add* anything in order to get the effect to occur. "[P]hysics is causally and explanatorily *self-sufficient*: there is no need to go outside the physical domain to find a cause, or a causal explanation, of a physical event" (Kim 2005: 16). But if that's true, according to this line of thought, then there simply is no room for their having any mental causes as well. Not, at least, if mental causes are distinct from physical causes.

The exclusion problem, then, results from the alleged fact that the three following theses cannot be jointly held:

*Mental-Physical Efficacy*: There are mental causes of physical effects

---

<sup>2</sup> This doesn't strictly follow. All that follows is that in order for the mental to bring about any changes in the world *something* must cause changes in the physical. But we will ignore this here. We find it highly plausible, in any case, that if your desire is to cause you to walk to the fridge it must do so by affecting some physical change in your body.

<sup>3</sup> One might reasonably question the truth of the closure principle (see e.g. Sturgeon 1998, Cartwright 1999) but we will grant it for the purposes of this discussion.

*Physical Causal Closure:* All physical effects have sufficient physical causes.

*Non-Reductivism:* Mental causes are distinct from physical causes.

This has been particularly forcefully argued by Jaegwon Kim in a body of work spanning more than two decades (see e.g. Kim 1989a, 1998, 2005). And the intended moral is that as nobody ought to reject efficacy or closure, the incompatibility shows non-reductivism to be untenable. Our aim here will be to defend non-reductivism against this charge.

Philosophers have responded to the exclusion problem in a number of different ways, some of which involve significant revisions of ontological and metaphysical assumptions. Some believe that the problem can only be solved by returning to reductive physicalism (e.g. Papineau 1990, Kim 1998) others that it should be solved by some kind of dualism (Lowe 2008); or that it needs a different account of the relata of causation (Robb 1997, Ehring 1999, Macdonald & Macdonald 1986; see also Gibb 2004); or that it needs a different account of causation (Menzies 2008, Raatikainen forthcoming). This is the form the debate has taken in the last decade or two. In this paper, we shall return to the general form of the mental causation problem and question one of the assumptions on which it rests.

In our view, we do not need to adopt a new metaphysics of causation, or of the causal relata, in order to defend non-reductivism against the exclusion problem. What we need is a proper demonstration of the assumptions underlying the problem. Moreover, in order to resist the problem in the way we suggest, we do not need to adopt any specific ontological views about the relata or metaphysics of causation, or on fundamental ontology in general. The correct metaphysics of causation is, of course, a huge and important topic, and a full understanding of mental causation obviously needs a full understanding of causation. But a response to the exclusion problem does not need this. To the extent, then, that the exclusion problem is *the* mental causation problem, the mental causation problem does not await an answer the question of what the correct fundamental ontology should be.

## **1. Whose Problem is it?**

Standardly, the exclusion problem is thought to be a problem for all and only non-reductive physicalists. But the problem arises neither for all nor for only non-reductive physicalists. In its current form, the exclusion problem originates from a

misguided criticism of Davidson's anomalous monism in the 1980s. Davidson's critics complained that although his theory made room for mental causes by identifying them with physical causes, it did not give an adequate causal role to the mental 'as mental'. (See e.g. Stoutland 1980, Honderich 1982, Johnston 1985). In essence, the complaint was that on Davidson's account all the work is done by the physical properties of events and this leaves no room for a causal contribution by mental properties.<sup>4</sup> This criticism of Davidson is based on a misunderstanding of his theory of events and causation, as a number of writers pointed out (e.g. Smith 1984, Crane 1992, McLaughlin 1993). For Davidson himself denies that properties play any role in causation. Causation, he holds, relates events as such, and does not hold 'in virtue of' the properties of those events, or in virtue of how they are described. We are here using the term 'non-reductivism' to be the thesis that mental causes cannot be identified with physical causes. But sometimes the term is used more narrowly, for the thesis that mental *properties* cannot be identified with physical *properties*. And if the term is used in this latter sense, then the exclusion problem does not arise for all non-reductivists. For Davidson is a non-reductivist in this sense and does not face the problem.<sup>5</sup>

The exclusion problem, as it is standardly put forward, is a problem for those who deny the identification of mental causes with physical causes, while accepting supervenience and closure. This includes, but is not restricted to, non-reductive physicalists. For non-reductivists may accept supervenience and closure without subscribing to physicalism. What makes a non-reductive view physicalist are two things: its commitment to the core physicalist thesis of global supervenience, and its

---

<sup>4</sup> More specifically, this is because causation implies law, but mental events can be seen as instances of laws only under their physical descriptions. But Davidson (1970) had argued that all mental events are physical events, based on a particular theory of causation – that causation is a relation between events and that events can be seen to instantiate laws only under some descriptions.

<sup>5</sup> It should be noted, perhaps, that although the exclusion problem misses its mark as a criticism of Davidson, it might nonetheless be thought to be a serious worry for anyone who holds that the (primary) causal relata are Davidsonian events, while wanting to give a causal role to mental properties. As the problem is standardly formulated, however, in terms of the sufficiency of one cause and the redundancy of another, it misses the mark against such views. There may be some alternative formulation of the problem, appealing to *in virtue of* locutions, on which a tension can be forced. But we shall not explore this here. It is our belief that if the standard formulation fails, any such weaker formulation will fail also.

commitment to the claim that any fundamental or ‘brute’ facts must be facts within physics (Horgan 1993). A view that accepts the former claim but rejects the latter deserves the name ‘emergentist’ rather than physicalist (see Crane 2001, cf. Kim 2010). Emergentists of this kind take the mental to be intimately dependent on the physical, but they take this to be a matter of brute fact, rather than a fact explicable by the facts of physics. Although emergentists need not accept the closure principle, they may well do so. And if they do, they too need a response to the exclusion problem.

## **2. What Exactly is the Problem?**

But what exactly is the exclusion problem? Supposedly, the truth of the closure principle precludes non-reductivists from saying that some physical effects have mental causes. But why should that be?

The closure principle does not entail that physical effects have no non-physical causes. For although the physical causes by themselves suffice to fix the physical effects, this does not rule out that certain physical effects (certain bodily movements, for instance) might have *both* physical and non-physical causes. Indeed, it is perfectly compatible with closure that certain physical effects (actions, for instance) *always* have both physical and non-physical causes. So non-reductivism, it would seem, is not in fact incompatible with efficacy and closure. To get from the claim that all physical effects have sufficient physical causes to the conclusion that physical effects have *only* physical causes, Kim relies crucially on two further assumptions:

*Denial of Overdetermination.* Mental causes do not overdetermine their effects.

*The Exclusion Principle:* “No single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of overdetermination.” (Kim, 2005, p.42)

It is against the backdrop of these assumptions, that non-reductivism can be seen to be incompatible with closure and efficacy. Indeed, if we combine these assumptions with closure and efficacy, we get a valid argument for the falsity of non-reductivism:

1. *Mental-Physical Efficacy:* There are mental causes of physical effects
2. *Physical Causal Closure:* All physical effects have sufficient physical causes.

3. *Denial of Overdetermination.* Mental causes do not overdetermine their effects.

4. *The Exclusion Principle:* “No single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of overdetermination.” (Kim, 2005, p.42)

Therefore,

C. *Identity of Causes:* Mental causes are identical to some physical causes.

The argument so formulated is neutral on what types of entities causes are. If you think causes are events, then you will take this as an argument for the identity of mental and physical events. If you think causes and effects are properties, you will take it as an argument for the identity of mental and physical properties. Kim’s own view is that instantiations of properties (or what he calls ‘events’) are causes, and he argues via the identity of such property instantiations to the identification of mental and physical properties (see e.g. Kim 1998). We will not assume any particular account of the causal relata in our discussion. We think the argument fails to establish the identity of causes, regardless.

A number of philosophers have recently rejected Kim’s exclusion argument (see especially Bennett 2003; Loewer 2007; Raatikainen forthcoming) by questioning some of the assumptions behind it. While these responses have their merits, they depend on adopting some more or less controversial theses in the philosophy of language or metaphysics (e.g. Lewis’s semantics for counterfactuals, causation as counterfactual dependence or causation as difference-making). Our approach, by contrast, is to undermine the argument by making the *weakest possible* assumptions about causation and ontology. We will argue that the exclusion principle is contrary to our ordinary judgements about causation, has no strong independent defence and ought to be rejected.

### **3. Overdetermination**

Before getting on to our criticism of the exclusion principle, we need to explain first what is meant by overdetermination and why mental causes and physical causes do not overdetermine their effects.

The cases most naturally thought of as cases of genuine causal overdetermination, are cases where there are two or more causes, each of which

would have been sufficient to bring about the effect *in the absence of the other*. A classic morbid example is death by two assassins working independently, where a man is killed by two bullets hitting him at once. Another is where a window is shattered by two rocks making impact with it at the very same time. What such cases have in common is that in whatever way the causes are sufficient, they are sufficient *independently of each other*. This is shown by the fact that had either of the causes been deleted, without anything being added in its place, the other cause would still have caused the effect. What makes the death of the assassinated man overdetermined, for instance, is that had one of the assassins not shot him, he would still have been shot dead by the other assassin. In his recent work, Kim has been explicit that in formulating the exclusion principle and denying overdetermination he has in mind overdetermination of this standard type.<sup>6</sup>

We should agree with Kim that mental and physical causes do not overdetermine their effects in this way. This is because mental and physical causes are not independent of one another. This leaves open, however, that mental and physical causes are nonetheless numerically distinct. For *distinct* sufficient causes need not be *independent* sufficient causes. Indeed, that there might be distinct but dependently sufficient causes is what we should expect if physicalism is true; and also if emergentism is true. Physicalists and emergentists hold that the mental is very *intimately dependent* on the physical, but identity is only one way in which this need be so. One popular way of spelling out this dependence is to say that mental properties are *realized* by physical properties, where realization may be understood as follows:

To realize is to ‘make real’ in a sense of ‘makes’ that is constitutive rather than causal. So a property-realizer of a property is a property whose instantiation constitutively makes real an instantiation of the realized property. (Shoemaker 2007: 10)

---

<sup>6</sup> “The usual notion of overdetermination involves two or more separate and independent causal chains intersecting at a common effect. Because of *Supervenience*, however, that is not the kind of situation we have here. In this sense, this is not a case of genuine causal overdetermination.” (Kim 2005: 48).

How exactly a theory of realization should be developed is a question that we need not address here. The important thing in the present context is just that in claiming that the relationship between mental properties and their physical realizers is *constitutive*, the physicalist requires a tighter relationship than mere correlation (even if this correlation is nomological). The relationship that holds between mental property instantiations and their physical realizers is not akin to that which holds between smoke and fire, for instance. It is more closely analogous to the sort of relationship that holds between a statue and the lump of clay that makes it up. (That, of course, is a relationship between particulars; but realization in the relevant sense is intended to be an analogous relationship between properties).

This analogy between realization for properties and constitution for particulars is very helpful in evaluating both the exclusion principle and the denial of overdetermination. Given this understanding of the relationship between the mental and the physical, mental causes and their physical realizers could not overdetermine their effects in the ‘independent assassins’ way. For holding everything else fixed, you could not delete one of them from a given context without thereby deleting the other. If such a relation were to hold between properties, then this is an excellent reason for denying that mental causes and their realizers overdetermine their effects.

The important thing to stress here is that even when combined with mental-physical efficacy and the closure principle, the denial of overdetermination gives us no obvious reason to reject non-reductivism. For an effect may have distinct causes without being overdetermined, granted that the causes in question are suitably dependent (see Mellor 1995: 103-4 for a similar point).

#### **4. The Exclusion Principle**

It is the fourth premise of the exclusion argument, the exclusion principle, that challenges the compatibility of non-identity with premises (1)–(3). According to this principle ‘no single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of overdetermination’. And if that’s true, then it is clear that non-reductivism cannot make room for mental causation, given the truth of closure and the denial of overdetermination. It is on this principle that the case against non-reductivism rests and the remainder of our discussion will be devoted to examining it.



#### 4.1. Causal Exclusion and Explanatory Exclusion

How did the exclusion principle enter this debate? In his first significant intervention in this debate, ‘Mechanism, Purpose and Explanatory Exclusion’ (1989), Kim talked in terms of what he called the principle of *explanatory* exclusion: ‘no event can be given more than one complete and independent explanation’ (1989: 79). But over the years, Kim has changed from talking about *explanatory* exclusion to talking about causation. His *causal* exclusion principle states that ‘[n]o single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of overdetermination’ (Kim, 2005: 42).

Kim does not say why he has moved from talk of explanation to talk of causation. But this move is certainly a move in the right direction, since the problem of mental causation is a problem about causation and not explanation.<sup>7</sup> There is not even a *prima facie* difficulty of mental explanations being incompatible with, or ‘crowded out by’ physical explanations (see Burge 1993). Any occurrence can be explained in countless ways, and there is no incompatibility between any physical explanation of an event and a mental explanation of the same event. So there seems to be little plausibility to the idea that one explanation ‘excludes’ another.

However, Kim’s ‘explanatory exclusion’ principle does talk about *complete* explanations. What is a complete explanation? Suppose it is the conjunction of *all* the many different true explanations of the event – if we can make sense of this idea. Then of course, no event can have more than one complete explanation in this sense. So it seems that the explanatory exclusion principle is either obviously false or trivially true. The same is not true of the *causal* exclusion principle. Though perhaps it is the connection with the (possibly illusory) idea of a ‘complete’ explanation of an event that lies behind Kim’s thought that the exclusion principle is an obviously true principle that requires no defense.

---

<sup>7</sup> It should be said that how closely related explanatory and causal exclusion are depends very much on your account of causation and of causal explanation. On Kim’s account the same things can serve as the relata of causation and causal explanation (i.e. property instantiations or facts) while for others, e.g. Davidson, causation and explanation relate very different things. It is only on the latter sort of view that causal explanatory exclusion is much less plausible than causal exclusion. It is fair to say, however, that if one wants to press a perfectly general problem of mental causation, one ought to avoid building in assumptions about the relation between causation and explanation. And on its face the principle of causal explanatory exclusion seems a lot less plausible than the principle of causal exclusion.

#### 4.2. *Why the Causal Exclusion Principle is Substantive*

Kim remarks that the causal exclusion principle is ‘virtually an analytic truth with not much content’ (2005: 51). But given what he means by ‘genuine overdetermination’, and given what he has to mean by ‘more than one’, this claim is clearly mistaken. There is a principle in the vicinity of the exclusion principle that is a better candidate for being ‘virtually an analytic truth’; namely, that no single event can have more than one *independently* sufficient cause occurring at any given time—unless it is a genuine case of overdetermination. That much just falls out of the notion of overdetermination under discussion. However, that principle entails no stronger conclusion than that mental causes are not sufficient *independently* of physical causes, and as we have seen that is not a claim that is in tension with non-reductive physicalism or emergence.

Kim’s causal exclusion principle is not specified in terms of *independently* sufficient causes, but only in terms of *distinct* sufficient causes. It says that there *cannot be* distinct sufficient causes of an event occurring at any given time except in cases of genuine overdetermination. This claim, we argue, is false. We have two main objections. First, it seems to us that far from being an analytic truth, the exclusion principle is not even plausible on its face. It conflicts with our causal judgments even before any physicalist commitments enter the picture and is subject to a number of counter-examples. Given this, we ought to demand some very good arguments to persuade us that the principle is true. Our second objection is that we lack such arguments.

#### 4.3. *On the Implausibility of the Exclusion Principle*

Notice, to begin with, that the principle as stated involves the important qualification ‘at the same time’. Without this qualification it would be refuted by the simple fact that every effect has many causes, stretching back across time. Think of a causal ‘chain’ where A causes B and B causes C. Even without assuming the transitivity of causation, A and B can both be causes of C. And if all causes are sufficient causes, then an effect can therefore have many sufficient causes across time: everyone should accept this. This is presumably why Kim adds the qualification, ‘at the same time’.

But why should it be more plausible with this qualification added? Even at a time, events have many causes, and not just in cases of genuine overdetermination.

J.L. Mackie's (1965) famous discussion of a short-circuit causing a fire illustrates this nicely. Putting to one side the distinction between cause and background conditions, it is natural to say that the presence of oxygen and the presence of flammable material are causes of the fire, just as much as the short-circuit is. But these are states of affairs, or property-instances, that exist at the same time as the fire. So the short-circuit can have many causes occurring at the same time.

But it may be said (as Mackie himself did) that none of these are *sufficient* causes. If 'sufficient' means *absolutely sufficient on its own*, then they are not sufficient causes. But is there anything which is absolutely sufficient on its own for the occurrence of an event? Those who believe in sufficiency in this sense might appeal to the fact that there is an entire state of the universe before the occurrence of the event which is sufficient for that event's occurrence. Now if determinism is true, then there must be such a state. Whether or not this gives any plausibility to the claim that no event can have more than one sufficient cause at the same time depends on the relationship between this idea of *the entire state of the universe*, and the idea of *something's being a cause*, or a sufficient cause. Of course, there have been theories which make a close connection between these ideas—Mill's notion of the 'whole cause' is the most famous—but these are specific accounts of causation, and are not uncontroversial.

By contrast, the exclusion argument is not supposed to rely on any particular account of causation. The exclusion principle is intended as a general principle that one ought to accept whatever one's account of causation and the causal relata: so we should expect the principle to accord with ordinary causal claims. But it does not appear to do so. Here is an example. We are in general happy to attribute causal powers to ordinary objects (see Lowe 2008). We say things like 'the furniture scratched the floor', 'my shoe gave me blisters' or 'the hammer made an indentation in the clay', for instance. But where we are happy to say that objects caused things we are often happy to say at the same time that their parts caused things. Suppose, for instance, that the indentation that my hammer makes in the soft clay on top of which I place it, is made by the hammer's head and not by its shaft. There seems no tension in saying both that the hammer caused the indentation in the clay and that its head did. But the hammer and its head are numerically distinct things, so this would violate the exclusion principle – so long as they are sufficient causes.

Are they sufficient causes? Not, of course, in the sense of ‘absolutely sufficient’—the sense in which Mill’s entire state of the universe is sufficient. But they are sufficient in the sense in which any cause we can *know about* is sufficient. The causes that we cite as sufficient or necessary are only sufficient or necessary *given* other factors, including other causes and maybe the laws of nature. This is what people mean when they sometimes say that causes are sufficient ‘in the circumstances’. Our everyday commonsense way of thinking about causes—part of the data which the metaphysics of causation is arguably supposed to explain—only requires causes to be sufficient in this sense.

#### 4.4. *On the Lack of Argument*

Given the above considerations, if we are going to defend the exclusion principle, we need a strong argument. Our second main claim is that we lack such an argument. Kim suggests in a number of places that it would be very *odd* if actions always had two sufficient causes.<sup>8</sup> But why is it odd? One answer is that it is odd if the causes are not just distinct but also *independent* and absolutely sufficient. It would, of course, be an astonishing coincidence if our actions always had two *independent and absolutely sufficient* causes, and this would cry out for some explanation. But the non-reductivist can deny both that the mental and physical case is a case of two absolutely sufficient causes, and that it is a case of independent causes.

First, non-reductivism can deny that the causes are absolutely sufficient. Physicalists will want to say, of course, that the physical causes are as close to absolute sufficiency as any cause gets. For example, they may say that given determinism and the transitivity of causation, the state of the universe at the Big Bang is a cause of today’s weather in Iceland. Of course, even this is only sufficient given the laws of nature, on the usual way of thinking about these matters. But nonetheless it is as close to absolute sufficiency as we get.

Second, the important point for non-reductivists is that the mental cause of a physical effect is not a candidate for being sufficient in anything like this absolute sense, given that it is dependent on its physical basis. Moreover, this dependency *explains* why the mental cause is not sufficient in the sense of being absolutely

---

<sup>8</sup> E.g. “[i]t is at best *extremely odd* to think that each and every bit of action we perform [has] two distinct sufficient causes” (1989: 86).

sufficient. So once we acknowledge this intimate relationship between the mental and the physical—assumed both by physicalists and emergentists—there is nothing strange about our actions having both mental and physical causes. It certainly wouldn't be a coincidence nor would it be something for which we lacked an explanation.

So what is the argument for the exclusion principle? Why should we accept that there cannot be distinct sufficient causes except for in cases of genuine overdetermination? It seems to us that the above considerations seriously undermine this contention. First, we seem only too happy to speak as if there were distinct sufficient causes where overdetermination clearly does not apply; i.e. where the causes in question are not independently sufficient. Second, once we acknowledge that the causes in question are *not independent* the reservations that one might have about allowing distinct sufficient causes ought to let up. Curiously, however, it turns out that Kim thinks it is precisely *because* the two causes are *not* independent that there is a problem in acknowledging both of them:

[O]ur problem is not exactly that of causal overdetermination, although both have to do with an overabundance of causes. It is important to see that the problem that we face arises *because* the two putative causes are *not* independent events. The difficulty is exactly that the causal status of the dependent event is threatened by the event on which it depends.  
(1998: 53)

But how exactly is the causal status of the dependent event threatened by the event on which it depends? Kim's answer seems to be that given that the causal powers of the dependent event are *determined* by the event on which it depends, it couldn't bring any *additional causal powers* to the picture, and this, he thinks means that it cannot play any causal role. (This is the basis of what he came to call in his 1998 book the 'supervenience argument'). In *Mind in a Physical World*, Kim claims 'there is a real problem, the exclusion problem, in recognizing second-order properties as causally efficacious *in addition to* their realizers' (1998: 53, our emphasis). He goes on to explain what the problem is, as follows:

For there is nothing in the instantiation of [the second-order property] *F* on this occasion over and above the instantiation of its realizer *H*. Given this, to think that this instance of *F* has causal powers *in excess* of these of *H* is tantamount to belief in magic. (1998: 54-55, our emphasis)

Again, the idea is that where an effect already has a sufficient cause, any simultaneous distinct sufficient cause would have to bring some *additional causal powers* to the scene. Let us call this the *motivating principle*.

We find this motivating principle no more plausible than the exclusion principle. We do not have to admit that instantiations of mental properties have *additional* causal powers (on any given occasion) to those bestowed on them by their realizers. Indeed, it is difficult to see how anybody who believes in global supervenience could think this possible. For if the thesis of global supervenience is true, then fixing the physical facts fixes all the facts there are, including causal facts. If we have to accept that this rules mental properties out from being causally efficacious, we still need to know why. Just stating that a thing could not be causally efficacious unless it brought additional causal powers to those already determined by the physical seems straightforwardly to *beg the question* against those who believe both that global supervenience holds and that there are causally efficacious things that are neither identical nor separate entities, but rather stand in the relation of constitution or realization to one another. In other words, the motivating principle begs the question against the very people against whom the exclusion argument is directed.

The comparison with particular objects is instructive at this point. Many philosophers hold that statues are *constituted by* rather than identical to the lumps of matter that make them up. Statues and lumps, they claim, are distinct material objects that share all their matter and microphysical properties. If this is right, then when you place a copper statue in a tub of water with the effect that the water level rises; you place numerically distinct material objects in the water; namely, the statue, and the lump of copper that constitutes it. What's more, it seems that the statue is sufficient (given the other causes) for raising the water level and also that the lump is sufficient (given the other causes) for raising the water level. But the statue and the lump certainly do not *overdetermine* the raising of the water level. Nor does the statue *add*

anything to water-raising powers already put in play by the lump. Those who believe that statues are distinct from lumps should take this as reason to reject Kim's principles.<sup>9</sup>

It might be replied, of course, that we should deny that statues are distinct from lumps: they are identical. But our present point is that we should not insist that this is so *simply* because at any given time the statue has no causal powers in addition to those of the lump. Because if statues *were* in fact constituted by lumps we should expect them to have no causal powers in addition to those of the lump at any given time. Note that the 'at any given time' clause is important here. There being a statue can be relevant in all sorts of ways. And on occasions where the lump does not constitute a statue, coming to do so might well give it new causal powers. But on any given occasion where the lump is such that it constitutes a statue, the statue brings no extra causal powers to those already determined by the lump.

Similarly, we should not argue that mental property instantiations must be identical to, rather than realized by, physical property instantiations on grounds that mental property instantiations don't have any causal powers in excess of their realizers. For if mental property instantiations were in fact realized by, rather than identical to, physical property instantiations, that's exactly what we should expect.

Our conclusion is that we have no good reason to accept the exclusion principle. It fits badly with ordinary causal judgments, it is unsupported by argument, and it begs the question against those who believe in constitutively related causes. Since non-reductivists—non-reductive physicalists and emergentists alike—agree that the causes in question are intimately related, there is no reason why they should be moved by the appeal to the exclusion principle.

## 5. Concluding Remarks

Once we have rejected the exclusion principle, we have disarmed the exclusion argument. A non-reductive physicalist or an emergentist can accept premises (1)–(3) of the argument, so long as they accept that mental and physical causes are intimately

---

<sup>9</sup> The example can also be reconstructed in terms of property instantiations. Given that property instantiations are individuated by the individuals instantiating the properties, the lump being submerged in the water and the statue being submerged in the water are distinct property instantiations, each of which is sufficient for the water level's rising; and neither of which adds any causal powers to those of the other one.

dependent. (It is a further question of whether this dependence can be explained—i.e. whether dependence is a ‘brute fact’ as some emergentists claim). Intimate dependency is also dependency of causal powers, and this is why non-reductivists can reject Kim’s motivating principle that non-identical causes would have to bring additional causal powers to those determined by the causes they depend on. These points are illustrated by the example of constitution between distinct objects.

What is more, we can disarm the exclusion argument without making very many heavy-duty assumptions about causation. The only substantive claims we are making about causation is that every effect can have many causes, both over time and at a time, and that these causes can be sufficient (in one of a number of senses typically appealed to by theories of causation). According to physicalists, closure entails that the physical causes are as absolutely sufficient for their effects as any cause can be. Mental causes will not be sufficient in this sense: they will only be sufficient given the other causes and other factors. But *contra* Kim, this is not something which undermines the mental cause’s status as a cause. For one thing, most of the causes we know about are sufficient ‘in the circumstances’ and not absolutely sufficient. But more importantly in this context, the mental cause is dependent on the physical cause. This is the essence of what it means to be a non-reductive physicalist, or an emergentist in our sense. So if these forms of non-reductivism are true, then we should not expect that mental causes would be absolutely sufficient.

If the closure principle is true, then the physical cause suffices for the physical effect. But since the mental cause is dependent on the physical cause, the latter also suffices for the former. Since the mental cause is, by hypothesis, a cause of the physical effect, this shows that the physical effect can have more than one cause. The mental cause is sufficient for this effect too, it’s just that it is sufficient given the physical cause. But this is the sense in which most causes are sufficient, or at least the sense in which most theories of causation allow themselves to talk of ‘sufficient causes’.

It would be a mistake to conclude from this that ‘all that matters’ is the physical cause. For this would assume that the physical cause could be there without the mental cause. But this contradicts the supposition, common to both non-reductive views being considered, that the physical cause necessitates the mental cause (see Loewer 2007, Loewer & LePore 1987.). So we should reject the idea that all that matters is the physical cause. Compare our analogy with the statue again: just because



the statue is determined by the arrangement of the clay, this does not mean that all that matters is that there is clay there. It matters that the clay gives rise to a statue.

Our response is available to those holding a wide variety of views about the relata of causation. Causes can be ordinary objects, or substances (cf. Lowe 2008). They can be events as conceived by Davidson (1967) or property-instantiations as conceived by Kim (1968) or differently conceived by Macdonald and Macdonald (1986). They can be *facta* in Mellor's (1995) sense. Or they can be tropes, as Ehring (1999) and Robb (1997) think. None of these views makes any difference to the way the exclusion problem should be treated. Once the causal exclusion principle is rejected, then it is clear that non-reductivism is not threatened by the conjunction of (1)—(3). The correct ontology can be argued about at a later stage.<sup>10</sup>

### References:

- Bennett, K. (2003). 'Why the Exclusion Problem Seems Intractable and How, Just Maybe, to Tract it'. *Nous*, 37(3): 471–97.
- Burge, T. (1993). 'Mind-Body Causation and Explanatory Practice'. In J. Heil and A. Mele (eds), *Mental Causation*. Oxford: Clarendon Press: 97–120.
- Cartwright, N. (1999). *The Dappled World*. Cambridge: Cambridge University Press.
- Crane, T. (1992). 'Mental Causation and Mental Reality'. *Proceedings of the Aristotelian Society*, 92: 1–18.
- (2001). *Elements of Mind*. Oxford: Oxford University Press.
- Davidson, D. (1967). 'Causal Relations'. Rpt. in (2001), *Essays on Actions and Events*, 2<sup>nd</sup> Ed. Oxford: Clarendon Press.
- (1970). 'Mental Events'. Rpt. in (2001), *Essays on Actions and Events*, 2<sup>nd</sup> Ed. Oxford: Clarendon Press.
- Ehring, D. (1999). 'Tropes in Seattle: The Cure for Insomnia'. *Analysis*, 59(261): 19–24.
- Garber, D. (1983). 'Mind, Body and the Laws of Nature in Descartes and Leibniz'. *Midwest Studies in Philosophy*, 8 (1):105–33.

---

<sup>10</sup> We would like to thank: the Leverhulme Trust and the Isaac Newton Trust for their generous support for Steinvör Árnadóttir's research; two anonymous reviewers for OUP for their helpful comments; and Jonathan Birch for his insightful remarks when some of this material was presented by Tim Crane in a seminar in Cambridge.

- Gibb, S.C. (2004). 'The Problem of Mental Causation and the Nature of Properties'. *Australasian Journal of Philosophy*, 82(3): 464–76.
- Guttenplan, S. (ed)(1994). *A Companion to Philosophy of Mind*. Oxford: Blackwell Publishers.
- Honderich, T. (1982). 'The Argument for Anomalous Monism'. *Analysis*, 42: 59–64.
- Horgan, T. (1993). 'From Supervenience to Superdupervenience: Meeting the Demands of the Material World'. *Mind*, 102: 555–86.
- Jackson, F. (1998). *From Metaphysics to Ethics*. Oxford: Clarendon Press.
- Johnston, M. (1985). 'Why Having a Mind Matters'. In E. LePore and B. McLaughlin (eds), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell.
- Kim, J. (1989). 'Mechanism, Purpose and Explanatory Exclusion'. *Philosophical Perspectives*, 3: 77-108.
- (1998). *Mind in a Physical World*. Cambridge MA: MIT Press.
- (2005). *Physicalism or Something Near Enough*. Princeton: Princeton University Press.
- (2010). 'Emergence: Core ideas and issues'. In *Essays in the Metaphysics of Mind*. Cambridge: Cambridge University Press.
- Leibniz, G.W. (1695). 'New system of the nature of substances'. In R. S. Woolhouse and R. Franks (trans and eds)(1998), *G.W. Leibniz: Philosophical Texts*. Oxford University Press.
- (1696). 'First explanation of the new system of the communication between substances' & 'Third explanation of the new system of the communication between substances'. In R. S. Woolhouse and R. Franks (trans and eds)(1998), *G.W. Leibniz: Philosophical Texts*. Oxford University Press.
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Basil Blackwell.
- (1994). 'Reduction of Mind'. In S. Guttenplan (ed): 412–31.
- Loewer B. (2007). 'Mental Causation, or Something Near Enough'. In B. McLaughlin and J. Cohen (eds), *Contemporary Debates in Philosophy of Mind*. Oxford: Basil Blackwell: 243–64.
- Loewer, B. & E. LePore (1987). 'Mind Matters'. *Journal of Philosophy*, 84: 630–42.
- Lowe, E. J. (2008). *Personal Agency: The Metaphysics of Mind and Action*. Oxford: Oxford University Press.

- Macdonald, C. & G. Macdonald (1986). 'Mental Causes and Explanation of Action'. *Philosophical Quarterly*, 36: 145–58.
- Mackie, J.L. (1965). 'Causes and Conditions'. *American Philosophical Quarterly*, 2(4): 245–64.
- McLaughlin, B.P. (1993). 'On Davidson's Response to the Charge of Epiphenomenalism'. In A. Mele and J. Heil (eds), *Mental Causation*. Oxford: Oxford University Press: 27–40.
- Mellor, D.H. (1995). *The Facts of Causation*. London/NY: Routledge.
- Menzies, P. (2008). 'The Exclusion Problem, the Determination Relation and Contrastive Causation'. In J. Hohwy & J. Kallestrup (eds), *Being Reduced: New Essays on Reduction, Explanation, and Causation*. New York: Oxford University Press: 196–217.
- Papineau, D. (1990). 'Why Supervenience?'. *Analysis*, 50 (2): 66–71.
- Raatikainen, P. (forthcoming). 'Causation, Exclusion, and the Special Sciences'. *Erkenntnis*.
- Robb, D. (1997). 'The Properties of Mental Causation'. *The Philosophical Quarterly*, 47: 178–94.
- Shoemaker, S. (2007). *Physical Realization*. Oxford: Oxford University Press.
- Smith, P. (1984). 'Anomalous Monism and Epiphenomenalism: A Reply to Honderich'. *Analysis*, 44(2): 83–6.
- Stoutland, F. (1980). 'Oblique Causation and Reasons for Action'. *Synthese*, 43: 351–67.
- Sturgeon, S. (1998). 'Physicalism and Overdetermination'. *Mind*, 107: 411–432.
- Woolhouse, R. (1985). 'Leibniz's Reaction to Cartesian Interaction'. *Proceedings of the Aristotelian Society*, 86: 69–82.